# ANALYSIS OF MODERATION PRACTICES IN A LARGE STEM-FOCUSED FACULTY

Iwona Czaplinski
Science and Engineering Faculty,
Queensland University of Technology,
Brisbane, Australia.

Lenore Adie
Faculty of Education,
School of Cultural and Professional Learning,
Queensland University of Technology,
Brisbane, Australia.

Bouchra Senadji
Science and Engineering Faculty,
Queensland University of Technology,
Brisbane, Australia.

Denise Beutel
Faculty of Education,
School of Cultural and Professional Learning,
Queensland University of Technology,
Brisbane, Australia.

.

***Abstract* — Moderation of assessment constitutes a crucial element of the learning and teaching process at the university. Yet, despite its importance, many academics have confusing beliefs and attitudes towards moderation practices, processes and procedures. This paper reports on a qualitative study conducted in a Science, Technology, Engineering and Mathematics (STEM)-focused faculty at a large Australian higher education institution. The findings of the study revealed a strong need for further investigation on the ways moderation is understood and enacted by academics within a STEM-specific context and informed redevelopment of the faculty's internal moderation policy.**

***Keywords: Moderation, Discourses of moderation, Assessment, STEM education***

## I. INTRODUCTION

This project undertook to investigate and describe the different internal moderation practices, processes and procedures used within a STEM-focused faculty in a large metropolitan Australian university. The overall goal of the research was to develop new internal moderation policy that would satisfy the requirements of a new national policy [1] demanding that details of moderation processes be made explicit for all assessment types and items. More precisely, the project aimed to ensure alignment of the faculty's quality assurance and management processes embedded in freshly developed courses with the recently reformulated Curriculum Quality Standards overseen by the Australian national university accreditation authority, the Tertiary Education Quality Standards Agency (TEQSA). These new arrangements aimed to improve accountability and transparency of Australian academic programs by embedding evidence-based practices of university degrees. Furthermore, the changed requirements also prompted discussion at the university level around multiple aspects of curriculum design, development, implementation and quality assurance with moderation procedures (internal and external) being one of the focal points of attention.

## II. LITERATURE REVIEW

Moderation remains a highly controversial problem in higher education with discussion focusing primarily around its purpose and accountability requirements for effective support of student learning and teaching [2]. According to Kuzich, Groves, O'Hare & Pelliccione [3, n.p.] there are two main purposes of moderation: (1) accountability that *enables an official confirmation of assessment quality* and (2) improvement that aims at establishing criteria for the quality of assessment by enabling markers to make *consistent and comparable judgments.* Sadler (2011) [4] argues in favor of a balanced approach towards both purposes required to ensure integrity of grades. While moderation practices, processes and procedures are regulated through policy in all Australian universities, its application varies widely in effectiveness, efficiency and consistency across the board. As identified in the literature [2], [5] the most often recurring problems surrounding moderation concern academic workload, limited choice of assessment types, feedback taking more time and insufficient improvement in reliability of standards. Bloxham [2, p.209] noted that not enough discussion on moderation in higher education takes place, which *allows assumptions of reliable standards to continue largely unchallenged*. It appears that over the past five years little has changed. This project attempted to address some of these issues by undertaking development of an evidence-based, research-led moderation policy for a large, STEM-focused faculty.

This research draws on work completed by Adie, Lloyd and Beutel [6, p. 968] and understands moderation as *a practice of engagement in which teaching team members develop a shared understanding of assessment requirements, standards and the evidence that demonstrates differing qualities of performance.*

At the university, moderation practices, processes and procedures are stipulated by policies and procedures that describing a broad set of principles that regulate activities and

operations related to various aspects of academic work. The policy differentiates between internal and external moderation practices, the first one ensuring cohesion of marks and grades awarded within particular units or subjects, while the second one operates at the course level (where the course is understood as a degree). External moderation establishes standards for professional accreditation bodies and warrants the reliability of assessment, grading and its consistency across higher education institutions by involving judgments by an external and independent expert.

During the scoping of the project, the authors identified additional areas requiring specific attention around the variety of moderation practices being employed by academics and the activities to be undertaken to improve them. As a result, the following specific objectives were established:

1. To gain knowledge of the range of internal moderation practices undertaken by academics;

2. To initiate work around the development of effective and efficient strategies for internal moderation practices, processes and procedures within the context of STEM-focused units to ensure the high quality of practices and impartiality of judgments.

In response to these objectives, the faculty surveyed all academics teaching at undergraduate and postgraduate levels for their internal moderation practices. While the findings of this study informed policy changes at faculty level it also revealed the need for further research and discussion in the area of moderation in Higher Education, specifically in the context of STEM-focused degrees.

## III. RESEARCH METHODOLOGY

This qualitative study comprised three phases: scoping, surveying and analysing the data in order to reach conclusions and lay foundations for the new policy.

The scoping phase focused on the design of the survey that sought to cover a broad spectrum of units, for instance at undergraduate and postgraduate levels, with large and small cohorts, units offered in all programs and courses, and the ones coordinated by academics working either independently or supervising activities of a large number of tutors. The survey also contained a range of seven current moderation practices identified through consultations with Academic Program Directors representing all five STEM-specific academic programs, staff discussions, informal enquiries and the authors' educational experience. The survey also comprised two open questions that enquired respectively about academics' perceptions of internal moderation practices and suggestions of possible improvements submitted by respondents.

In the second phase an online questionnaire was set up using 'Survey Monkey', a free online software and questionnaire tool. In total, 238 academics employed full-time and holding permanent positions were invited to express their opinion by filling in an online questionnaire. The survey was opened for 5 working days, from 21 to 26 August 2013, and registered 47 responses, providing rich data for analysis.

The table below summarises the design of the survey.

TABLE I.

| |
|---|
| **Section I** |
| 1. Number of students enrolled in your unit/s. If you are coordinating more than one unit, please list all numbers and separate them by a comma (e.g. 67, 259, 31).<br>2. Number of markers in your unit/s? If you are coordinating more than one marker, please list all numbers and separate them by a comma (e.g. 1,3, 10). |
| **Section II** |
| 3. Below are examples of moderation practices/ strategies currently used in the Faculty. Please indicate if, in your moderation practice, you are using one, some or all of them:<br>a) Establish precise requirements through Criterion-Referenced Assessment (CRA).<br>b) Minimise number of markers for a given assessment.<br>c) Use a single marker per question or a specific aspect of the assessment.<br>d) In the case of a single marker, moderate by having samples of marked assessments reviewed.<br>e) If the same piece of assessment is marked by more than one marker, combine the grades achieved appropriately (for example through averaging).<br>f) In the case of more than one marker, organise moderation meetings with your tutors/ markers to review samples of marked assessments and evaluate them against CRA sheets.<br>g) If there is more than one marker, organise moderation meetings with your tutors/ markers to compare assessed papers for consistency (for example through rotating samples of marked papers/ reports through the various tutors/ markers).<br><br>Other strategies |
| **Section III** |
| 4. What are positive aspects of your moderation practice?<br>5. What are negative aspects of your moderation practice? |

The third phase of the project allowed data analysis, and informed the development of new internal moderation policy.

## IV. DATA ANALYSIS

A relatively high response rate (19%) might suggest that overall moderation constitutes an important element of learning and teaching practice for STEM academics.

The data showed a variety with regards to the number of students enrolled in the unit, extending from a minimum of six to the maximum of 1060. As for the markers, their number spread from one to 48, averaging between three to five markers by unit. It is important to note that the data cannot reveal the correlation between the preferred moderation strategy and the size and/ or the level of the unit.

The responses of the second section of the survey provided evidence of the underlying concern of respondents for clarity, consistency, fairness and collaboration when it comes to moderating students' assessments. Of the respondents, 61% indicated that their primary moderation practice focused on establishing precise requirements through Criterion-Referenced Assessment (CRA). Other responses indicated different strategies for consistency of marking achieved through synchronisation of collaboration within the team and assurance that each marker was held responsible for a specific part of the assigned task. For instance, of the respondents 44% identified using a single marker for a given assessment, 34% declared minimising number of markers for a given assessment, 29%

said that they would review samples of marked assessment in the case of a single marker, and 23% would organise moderation meetings with their tutors and/ or markers to review samples of marked assessments and evaluate them against CRA sheets. 19% indicated that, if /when there are multiple markers, they would organise moderation meetings with the team in order to compare assessed papers for consistency. Finally, 12% of participants declared using a grade averaging technique in the case of more than one marker assessing the same piece of assessment.

The table below summarises the responses.

TABLE II.

| Moderation practice(s) | % of responses |
|---|---|
| Establish precise requirements through Criterion- Referenced Assessment (CRA). | 61% |
| Minimise number of markers for a given assessment. | 34% |
| Use a single marker per question or a specific aspect of the assessment. | 44% |
| In the case of a single marker, moderate by having samples of marked assessments reviewed. | 29% |
| If the same piece of assessment is marked by more than one marker, combine the grades achieved appropriately (for example through averaging). | 12% |
| In the case of more than one marker, organise moderation meetings with your tutors/ markers to review samples of marked assessments and evaluate them against CRA sheets. | 23% |
| If there is more than one marker, organise moderation meetings with your tutors/ markers to compare assessed papers for consistency (for example through rotating samples of marked papers/ reports through the various tutors/ markers). | 19% |

This might suggest that STEM academics perceive averaging as a 'mechanical' process that ignores the importance of the inherent educational principles focused at achievement of specific learning outcomes. We hypothesise that for the above-mentioned reason, this technique was rated the least favourable moderation strategy.

The open comments provided further examples of internal moderation strategies employed by the academics. Based on this information, we established a list of the additional moderation practices:

1. *Marking collectively:* all assessments are marked at the same time during a marking meeting.
2. *Checking for consistency* (from the position of authority): with multiple markers, at the beginning of the marking process, the unit coordinator moderates marks assigned by each individual marker.
3. *Calibrating marking at pre and post marking meetings:* team members mark randomly selected assessment items with special attention paid to consistency. This procedure is repeated at the end of the marking period.
4. *Providing exemplars:* Unit coordinator provides exemplars of previously marked assessments, including CRAs, marked at high, medium and low levels and accompanied with sample questions asked by students.
5. *Developing marking scheme:* unit coordinator provides marking team with a well-defined marking scheme.
6. *Focusing on feedback to students on marked piece of assessment*.

The addition of these moderation strategies by academics indicates their willingness to emphasise the practices that inform the collaborative aspect of their moderation models while providing a reference point in the form of a set of assessment standards (e.g. model 5). It is important to note the inclusive nature of these models embracing both, marking team members and students.

## V. DISCUSSION

One of specific objectives of the research was to gain knowledge of the range of internal moderation practices employed in the STEM-focused faculty of a large, metropolitan Australian university. Based on data analysis, two findings emerged regarding the ways moderation is understood and enacted in the faculty.

The first finding concerns the nature of the moderation practices. It appeared that moderation practices employed in the faculty diverged in the operational and logistical aspects (i.e. the number of meetings, who marks which part of the assessment, how the discussion of the assessed papers is organised), but converged from the perspective of quality assurance. More precisely, there was no one, common understanding of the concept of moderation across the faculty. Instead, respondents seemed to share a set of convictions about constituent elements of this notion. The findings indicated that the participants broadly understood the concept of moderation as a coordinated effort of all marking team members to ensure clarity, consistency, fairness and collaboration when assessing students' work and providing them with feedback.

Amongst key descriptors used by respondents in open comments, the term 'consistency' was used directly and appeared 4 times in 12 statements. One respondent commented: *Ranges and averages for each marker are checked to ensure a level of consistency is applied.* When reflecting on alternative strategies of moderation used by academics in their practice, another participant stated: *Moderation is to be applied if grading is inconsistent*. Interestingly, the latest comment indicates that, within this particular context, moderation is being used only in the case of inconsistencies between marks awarded by multiple markers. This might suggest that, at least some academics, associated moderation with standardising activity, making sure that all markers assessed students' work in the same way. If that is the case, important questions regarding, on one hand the quality of

educational expertise of markers and on the other hand the quality of students' learning experience should be asked.

While clarity and fairness did not appear directly in the survey, being precise and treating everyone equally can be extracted from the responses. For instance, the following comment was used by one of the unit coordinators to describe an alternative moderation strategy: *As the number of markers is small, I distribute examples of marked assignments to the markers in different grades, and I review papers, especially borderline cases as they are marked.* Another respondent wrote: *All markers work in their own field of expertise, so no moderation is required there. But we always meet after the results are collated to agree on the results, and always discuss any specific student who had non-academic difficulties. We agree on final grade cut-offs and special consideration outcomes.*

The collaborative nature of moderation procedures was made noticeable through such statements as: *Assessments are marked all at the same time in a marking party*, or: *We will be having two moderation meetings. The first occurs at the beginning where each tutor marks 4 or 5 randomly selected assignments and compare for consistency. A second meeting is scheduled once all papers have been marked, and a different selection of random assignments are used to check the consistency*. Note that, the last quotation also directly refers to consistency.

The first finding supports the observation that the concept of moderation is broadly understood as an amalgam of constituent notions with clarity, consistency, fairness and collaboration being the most prominent. Literature indicates that such understanding of the concept of moderation is not unique to STEM academics, but can be observed across various disciplines. A recent study by Adie, Lloyd and Beutel [6] on moderation practices employed in the Faculty of Education at the same university established a taxonomy of discourses impacting on academics' understanding of moderation. These are: (1) equity, (2) justification, (3) community building, and (4) accountability. When comparing the taxonomy with the findings of this study, we conclude that, while there are some important similarities in the ways academics from both faculties perceive, understand and act on moderation, there are also some important differences. For instance, the justification discourse seemed to be less strongly represented in the responses provided by STEM academics. This type of discourse focuses on *confidence in decisions, providing quality feedback, and support to respond to student queries* [6, p.973]. There was little evidence of explicit concern expressed by academics in relation to the quality of feedback provided to students. However, the quality of feedback directly impacts on student achievement of learning outcomes. A possible explanation of that omission might be found in the construction of the survey itself which did not provide any prompts in relation to above-mentioned aspects. On the other hand, being an inherent part of the moderation, they should be included *per se* in the discussion around this process and, more broadly, assessment. If one of the core aims of the moderation is to ensure the quality of learning and teaching experience, such silence speaks loudly about the need of further research with special attention paid to pedagogical and contextual aspects influencing the enactment of moderation. Questions investigating the level of markers' pedagogical preparedness, content expertise, marking experience, the level of communication between marking team members, assistance provided by the unit coordinator to all markers, or the presence/ absence of dialogue between all participants of the learning and teaching process (i.e. students, unit coordinator, markers) should be asked.

The second finding from the study also relates to learning and teaching process. Data analysis led us to hypothesise that academics were using Criterion-Referenced Assessment tool in a mono-dimensional way that is only as a benchmarking/ standardising tool. While the majority (61%) of respondents stated that their primary moderation strategy was to establish precise requirements through CRA, there was limited reference to CRA in open comments. The term 'CRA' was mentioned only once (sic) by one respondent who wrote: *Provide sample of CRAs of previous assignments marked at low, medium and high levels, students ask assessment questions of the tutor marking their assessment.* These comments indicate that the respondents perceived CRA as a suitable instrument to award grades and, in consequence, put more emphasis on the benchmarking/ standardising aspect of this evaluation tool than on its role of judging the quality of students' work by identifying and describing their level of achievement of learning outcomes.

A possible explanation of such an observation might be found in the very nature of being a STEM academic. It could be hypothesised that respondents translated CRA as a standardisation tool ensuring clarity, consistency and fairness when judging students' work. The fact that respondents were coming from STEM-focused disciplines, heavily depending on standards, set criteria, benchmarks and other precise measuring tools, might play an important role in the ways these academics perceived and understood the concept of assessment. In short, it appears that the nature of the discipline influenced the ways the moderation is understood and enacted. This observation needs further investigation.

## VI. CONCLUSION

This study has presented the results of a project attempting to unravel the different ways the internal moderation practices, processes and procedures are understood and enacted by academics at a STEM-focused faculty in a large Australian metropolitan university. We undertook this project in response to changes to the national policy [1] recently put in place. This prompted work around curriculum re-design and re-development, which encompassed re-writing policies that would comply with new standards and requirements.

Through a qualitative research and data analysis, we arrived at two main findings. First, there was no evidence that surveyed academics had a common understanding of the concept of moderation. Instead, the academics appeared to share a set of notions constituting the term, with clarity, consistency, fairness and collaboration being the most prominent. These coincided with taxonomy of moderation discourses identified in a previous study conducted in a Faculty of Education. Such observation allowed us to hypothesise that, although there are some similarities in the ways academics

from all disciplines understand the concept of moderation, there might be some discipline-specific differences. In short, we suspect that STEM academics might have a different, discipline-enhanced understanding of moderation. Further investigation is needed to either confirm or reject the above-mentioned hypothesis.

Second, we observed that some STEM academics might perceive moderation, and more broadly assessment, as an equity-centred activity and for this reason would have the tendency of using CRA as a benchmarking/ standardisation tool. We also noticed academics' silence regarding the achievement of learning outcomes by students. This raises more questions with regards to the ways STEM-academics perceive and understand the role moderation plays in the assessment, or more broadly, in the learning-teaching process.

All the above led us to identify further research questions investigating academics' beliefs, attitudes and the ways moderation is employed within a STEM-specific environment at university level. We also anticipate issues arising from implementation of freshly developed internal moderation policy. The implications of our findings for the STEM faculty's moderation policy revolve around the notion of moderation as an ongoing and integral component within the teaching-learning cycle. Support for STEM academics to view moderation in this manner needs to acknowledge their discipline-specific knowledges and ways of working and progress from this point, rather than introduce new ideas that may not be linked to their conceptual orientation.

## REFERENCES


[1] Tertiary Education Quality Standards Authority (TEQSA). (2012). Application Guide: Application for Accreditation of a Higher education course of study (AQF Qualification). Commonwealth of Australia.

[2] Bloxham, S. (2009). "Marking and moderation in the UK: False assumptions and wasted resources". Assessment and Evaluation in Higher Education, 34(2), 209-220.

[3] Kuzich, S., Groves, R., O'Hare, S. and Pelliccione, L. (2010). "Building team capacity: sustaining quality in assessment and moderation practices in a fully online unit". Paper presented at ATN Assessment Conference, 18-19 November 2010, University of technology, Sydney, Australia. Retrieved 18 September 2014 from http://www.uts.edu.au/partners-and-community/conferences-and-seminars/atn-assessment-conference/authors-and-presentations

[4] Sadler, D., R. (2011). "Academic freedom, achievement standards and professional identity". Quality in Higher Education, 17(1), 85-100.

[5] Elliot, E., Pearce, K., King, S. (2011). "Moderation of assessment tasks: Developing solutions to common problems". Paper presented at the 2011 HERDSA Conference, Gold Coast, Australia, July.

[6] Adie, L, Lloyd, M., Beutel, D. (2013). "Identifying discourses of moderation in higher education". Assessment and Evaluation in Higher Education, 38(8), 968-977.

[7] Adie, L. E. (2013). "Moderation practices in a faculty of education : the four discourse model". In Proceedings of the Australian Teacher Education Association Conference, Queensland University of Technology, Brisbane, QLD, pp. 1-6. Retrieved 18 September 2014 from http://www.atea.edu.au/ConfPapers/2013%20Peer%20Reviewed/Paper3_LenoreAdie.pdf

[8] Metha, A. and Xavier, R. (2007/8). "Building self-evaluation skills through criterion-referenced assessment in public relations". PRism 5(1&2). Retrieved 18 September 2014 from http://praxis.massey.ac.nz/prism_on-line_journ.html

[9] Rust, C., O'Donovan, B., Price, M. (2005). "A social constructivist assessment process model: How the research literature shows us this could be best practice". Assessment and Evaluation in Higher Education, 30(3), 231-240.